

BIO392-cnv-freq

Yasmine Douar

28.09.2022

Step 1: Install package

```
if (!require(devtools)){ install.packages('devtools')}

## Loading required package: devtools
## Loading required package: usethis
if (!require(pgxRpi)){ devtools::install_github('progenetix/pgxRpi')}

## Loading required package: pgxRpi
library(pgxRpi)
```

Step2: Search esophageal adenocarcinoma NCIt code

C4025

Step3: Access the CNV frequency data from samples with esophageal adenocarcinoma

```
frequency <- pgxLoader(type = "frequency", output = "pgxseg", filters = "NCIT:C3058", codematches=T)
```

The retrieved data is an object containing two slots meta and data.

The meta slot looks like this:

```
frequency$meta

##           code           label sample_count
## 1 NCIT:C3058 Glioblastoma         4384
## 2          total                    4384
```

The data slot includes two matrices.

```
names(frequency$data)

## [1] "NCIT:C3058" "total"

The frequency matrix looks like this
head(frequency$data$`NCIT:C3058`)
```

```
##           filters reference_name  start    end gain_frequency loss_frequency no
## 1 NCIT:C3058           1         0 400000         4.174         4.151  1
```

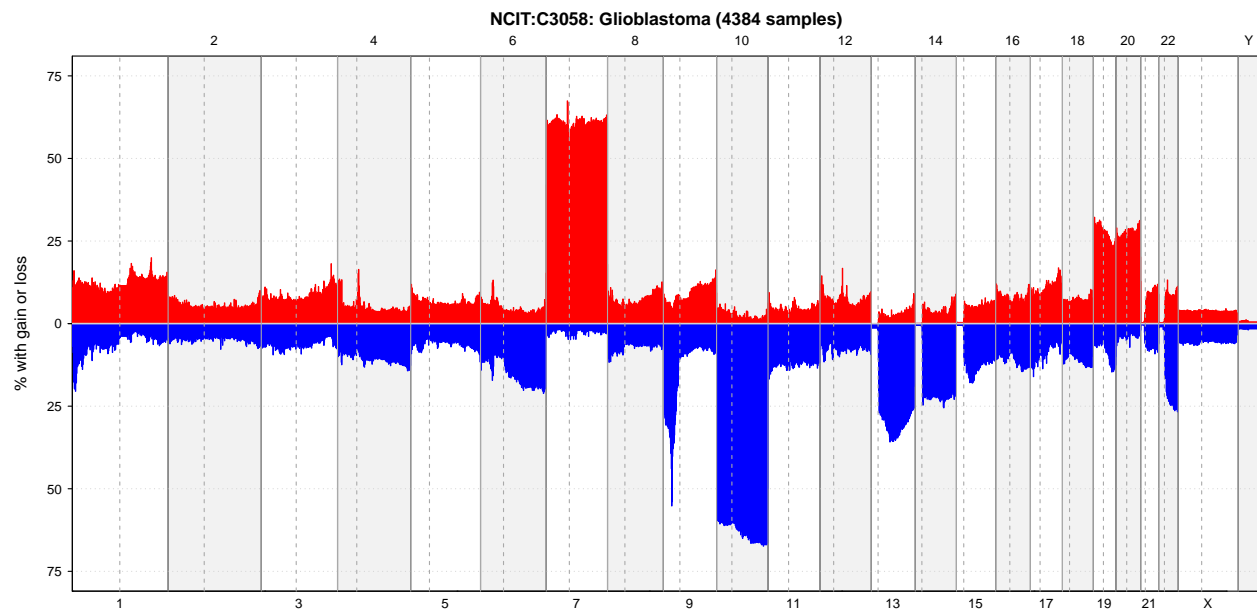
##	2	NCIT:C3058	1	400000	1400000	10.903	6.911	2
##	3	NCIT:C3058	1	1400000	2400000	13.047	8.805	3
##	4	NCIT:C3058	1	2400000	3400000	14.690	15.420	4
##	5	NCIT:C3058	1	3400000	4400000	16.036	17.769	5
##	6	NCIT:C3058	1	4400000	5400000	11.747	19.662	6

Dimension of this matrix

Step4: Visualize data

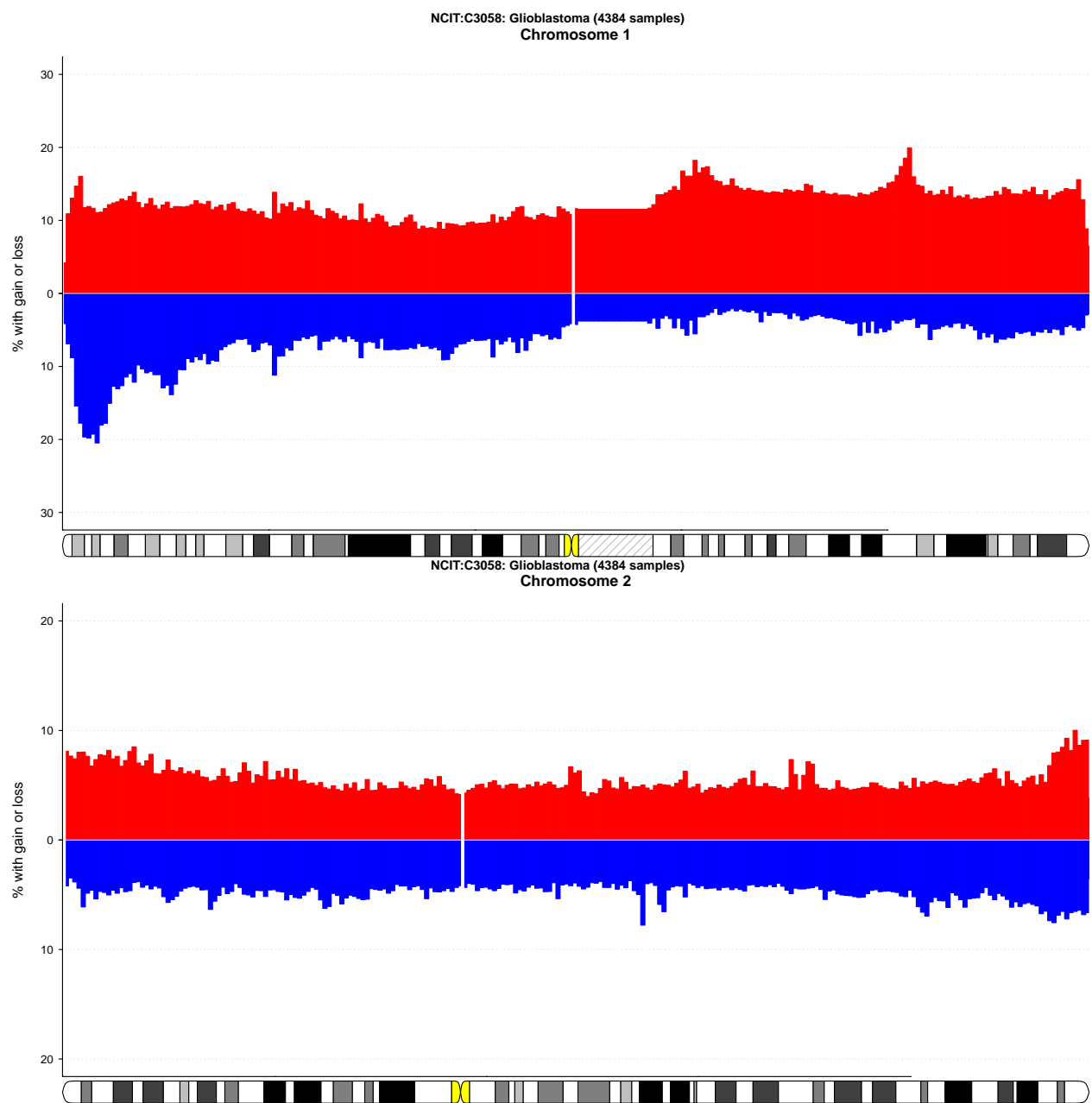
By genome

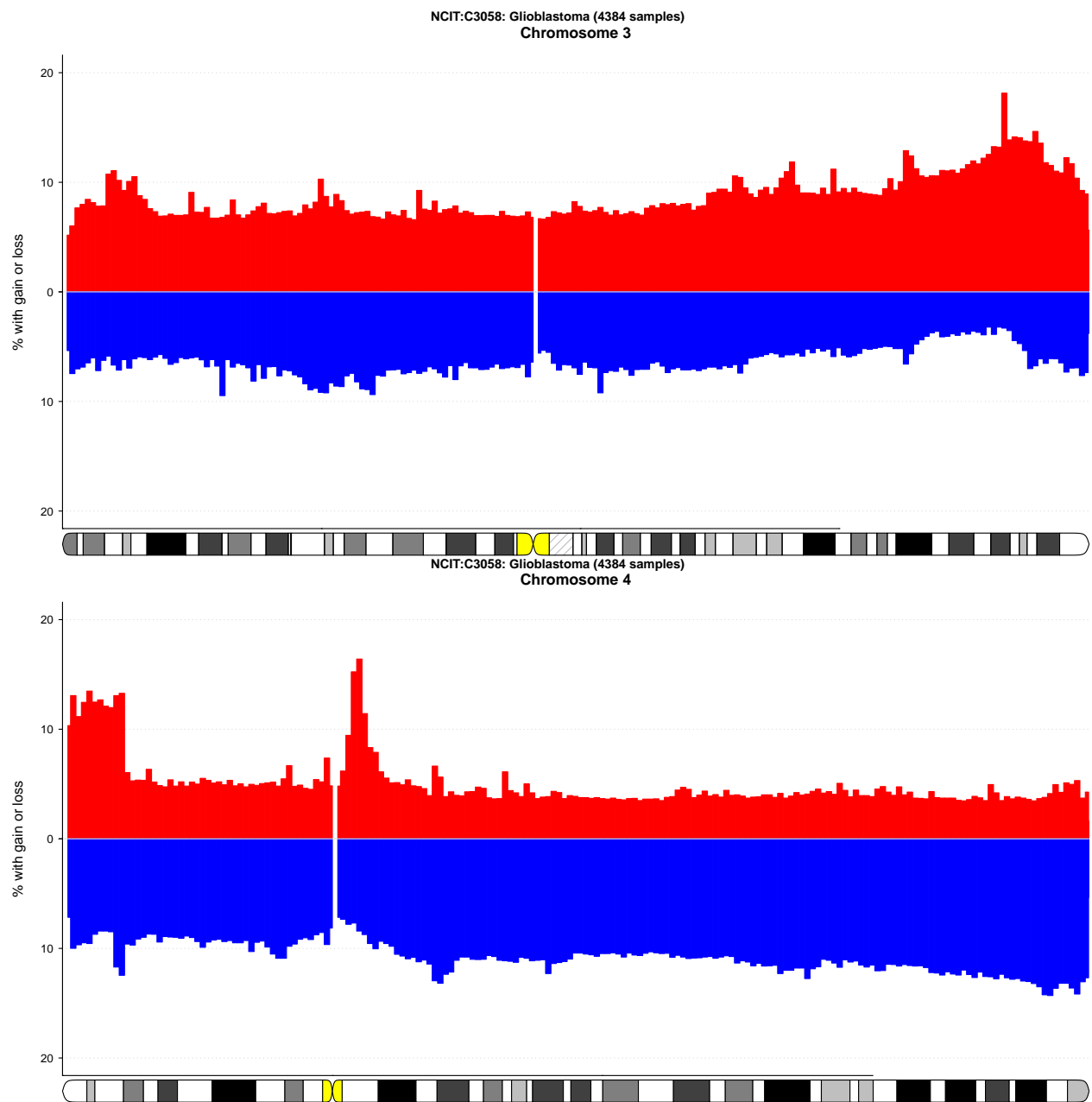
```
pgxFreqplot(frequency)
```



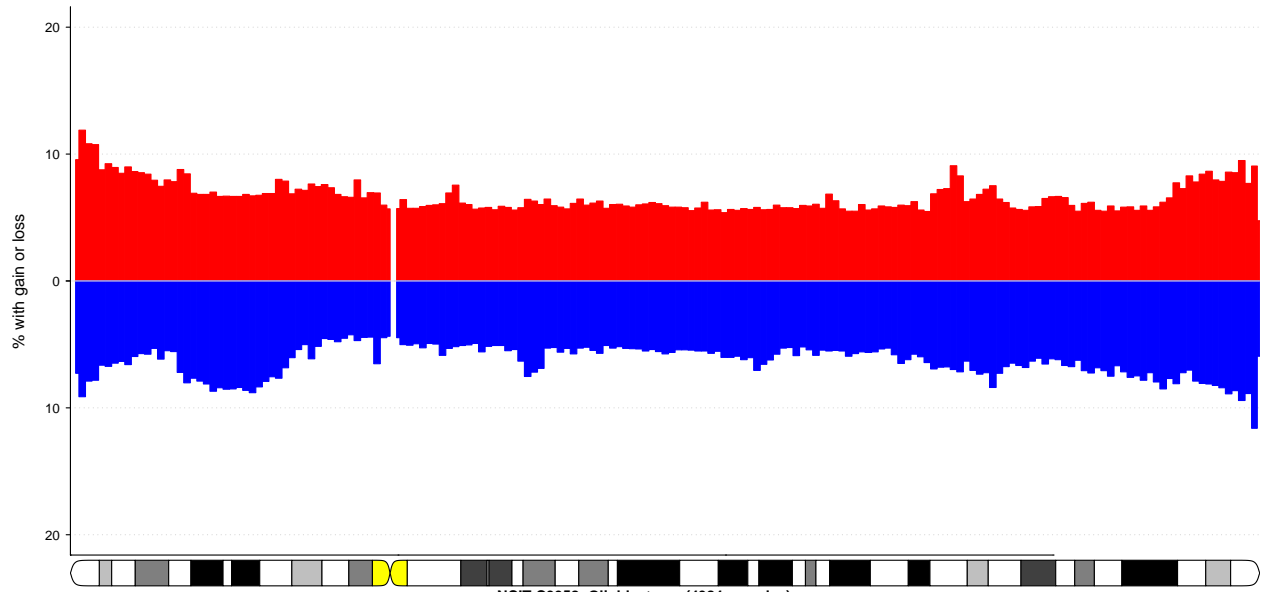
By chromosome

```
for (i in 1:24){
  pgxFreqplot(frequency,chrom = i)
}
```

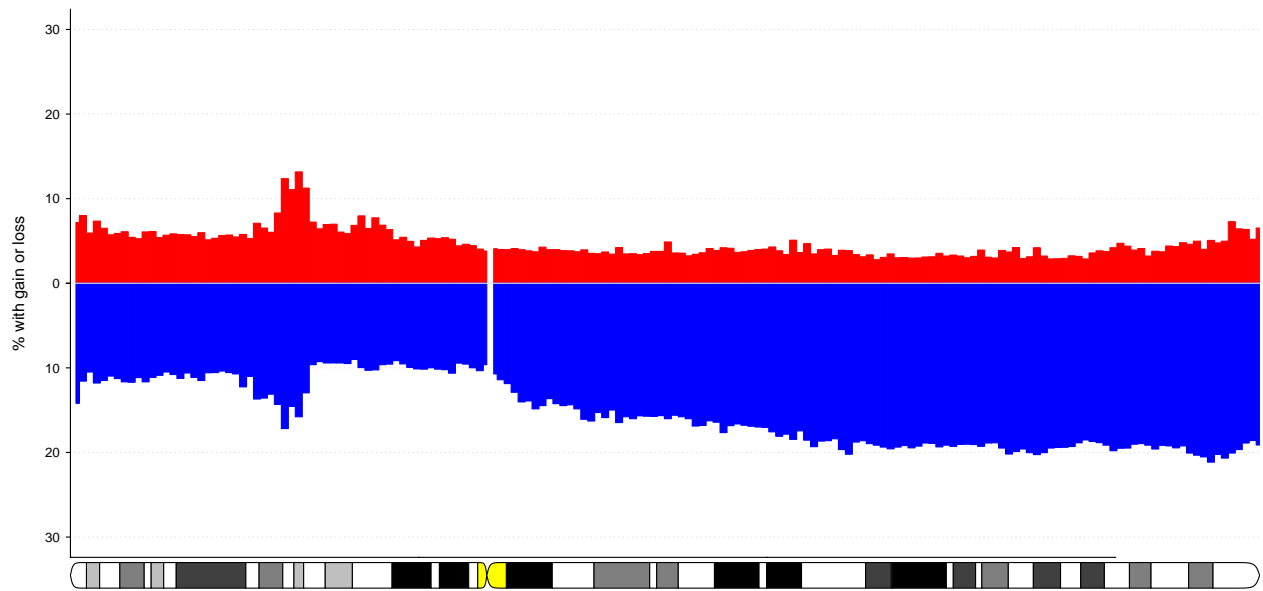


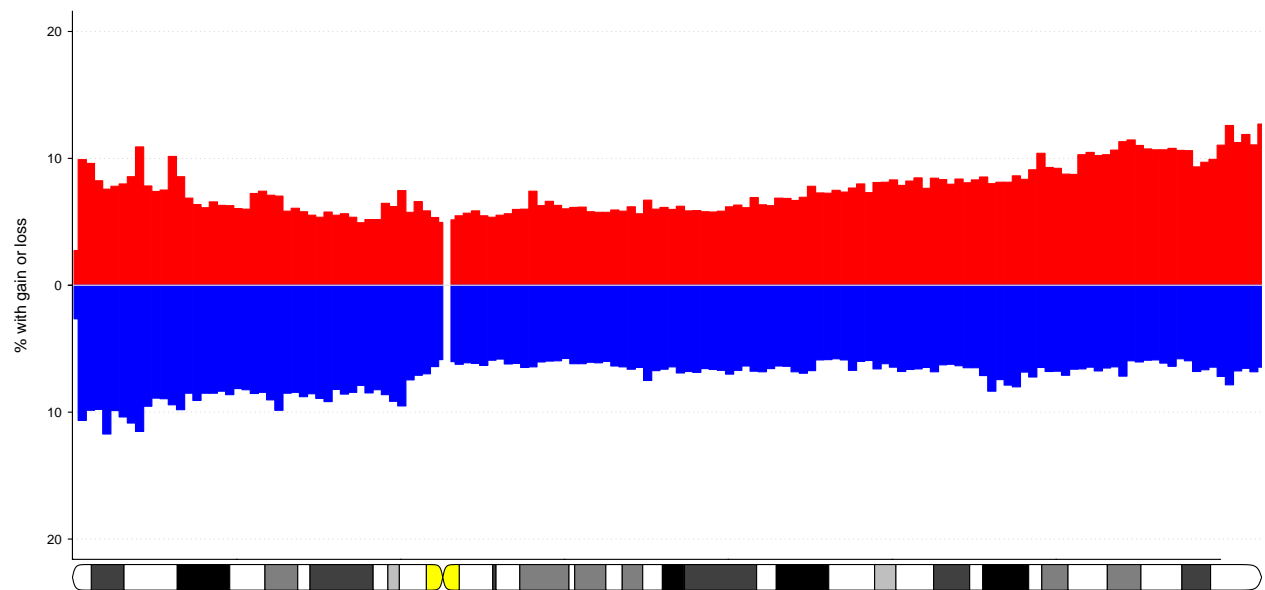
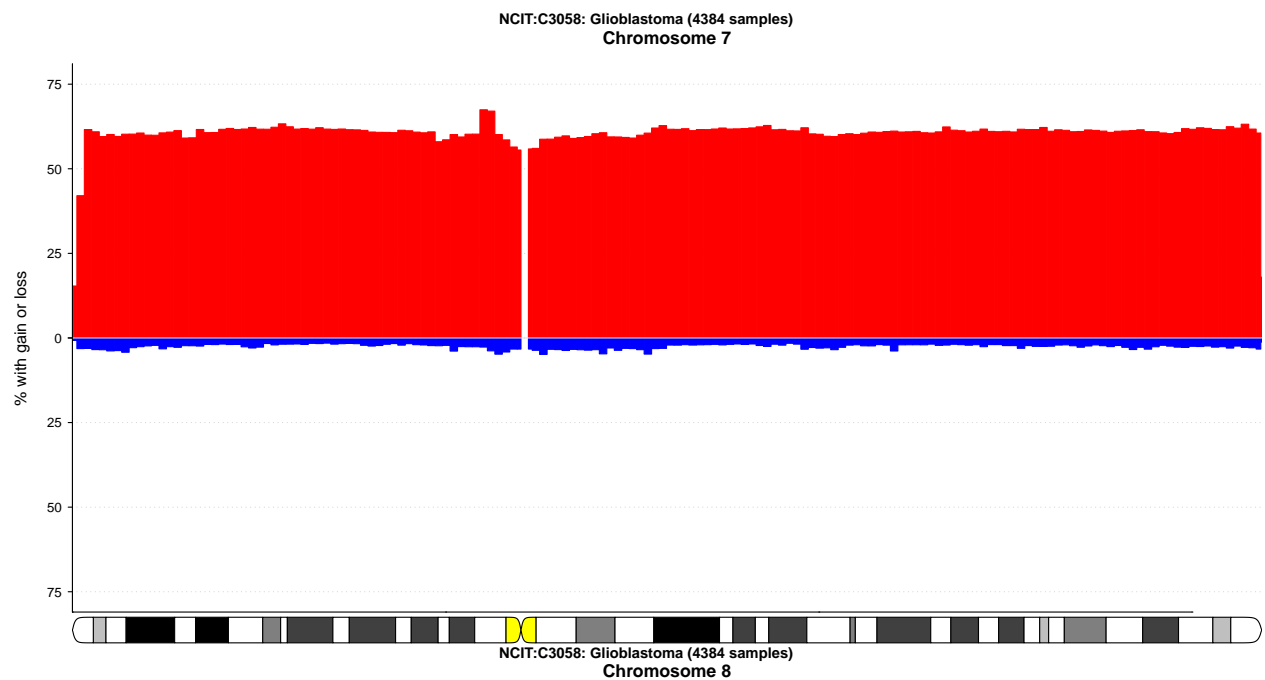


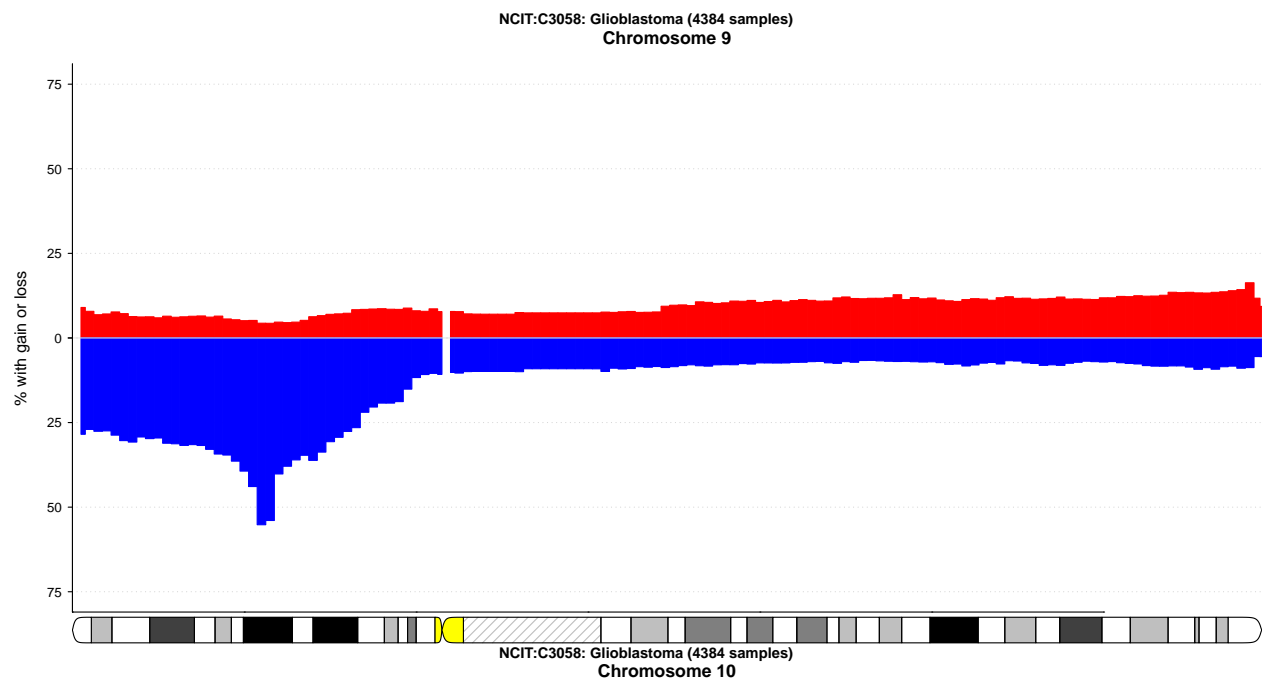
NCIT:C3058: Glioblastoma (4384 samples)
Chromosome 5

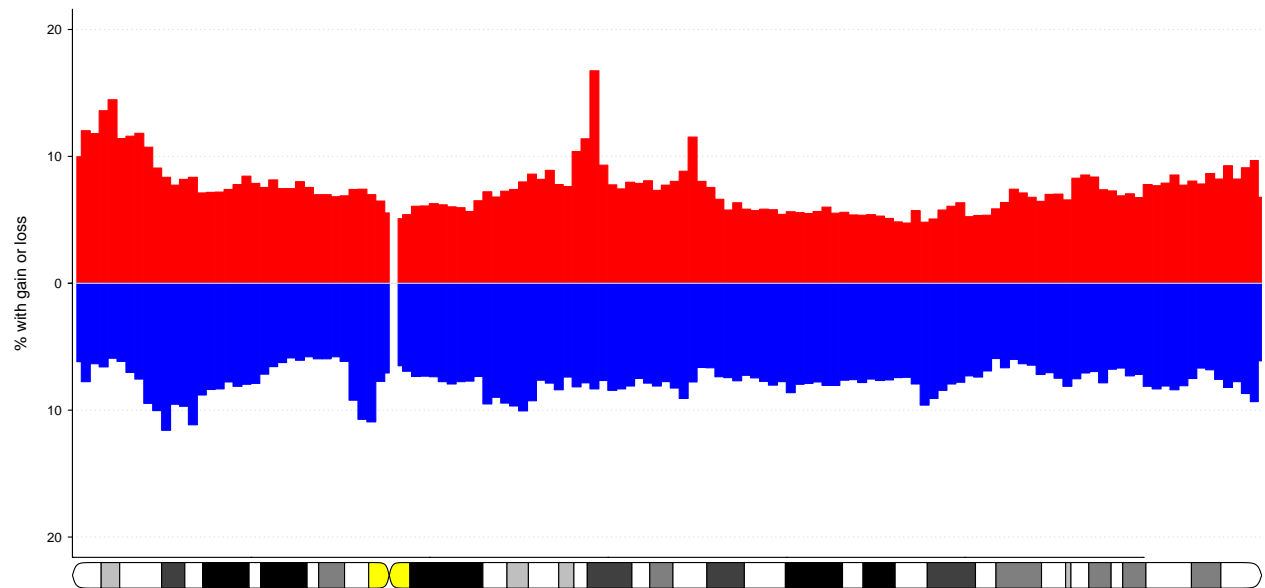
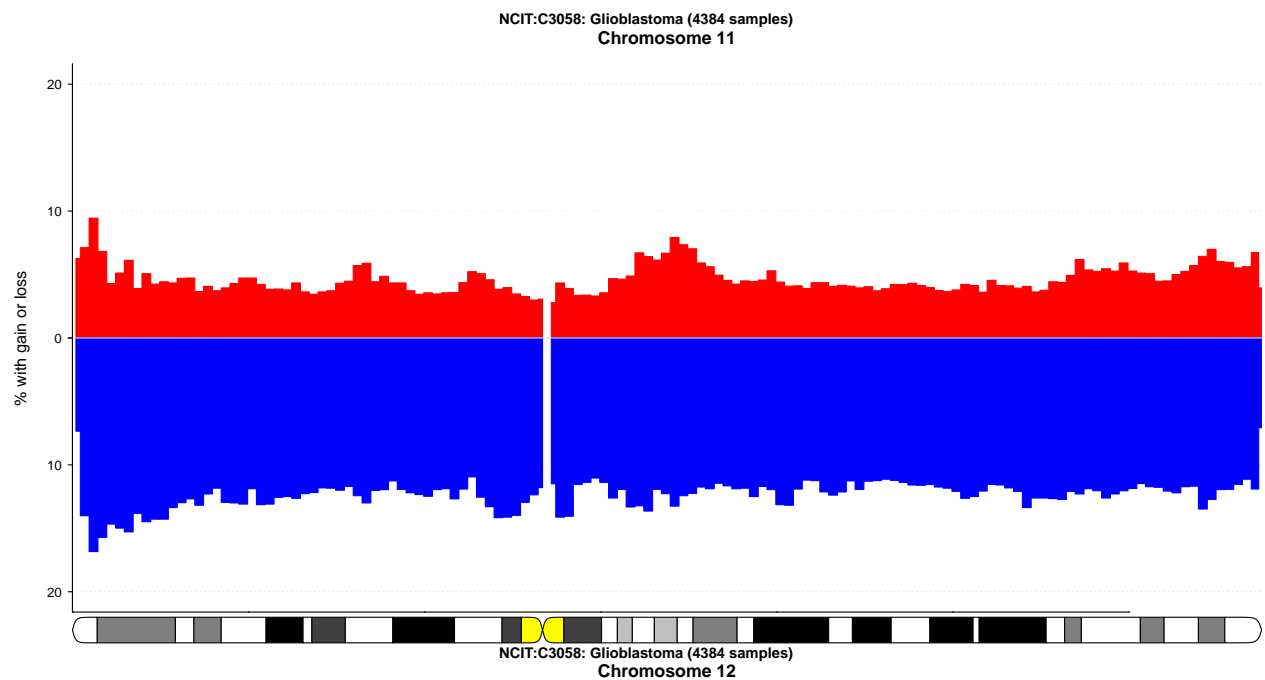


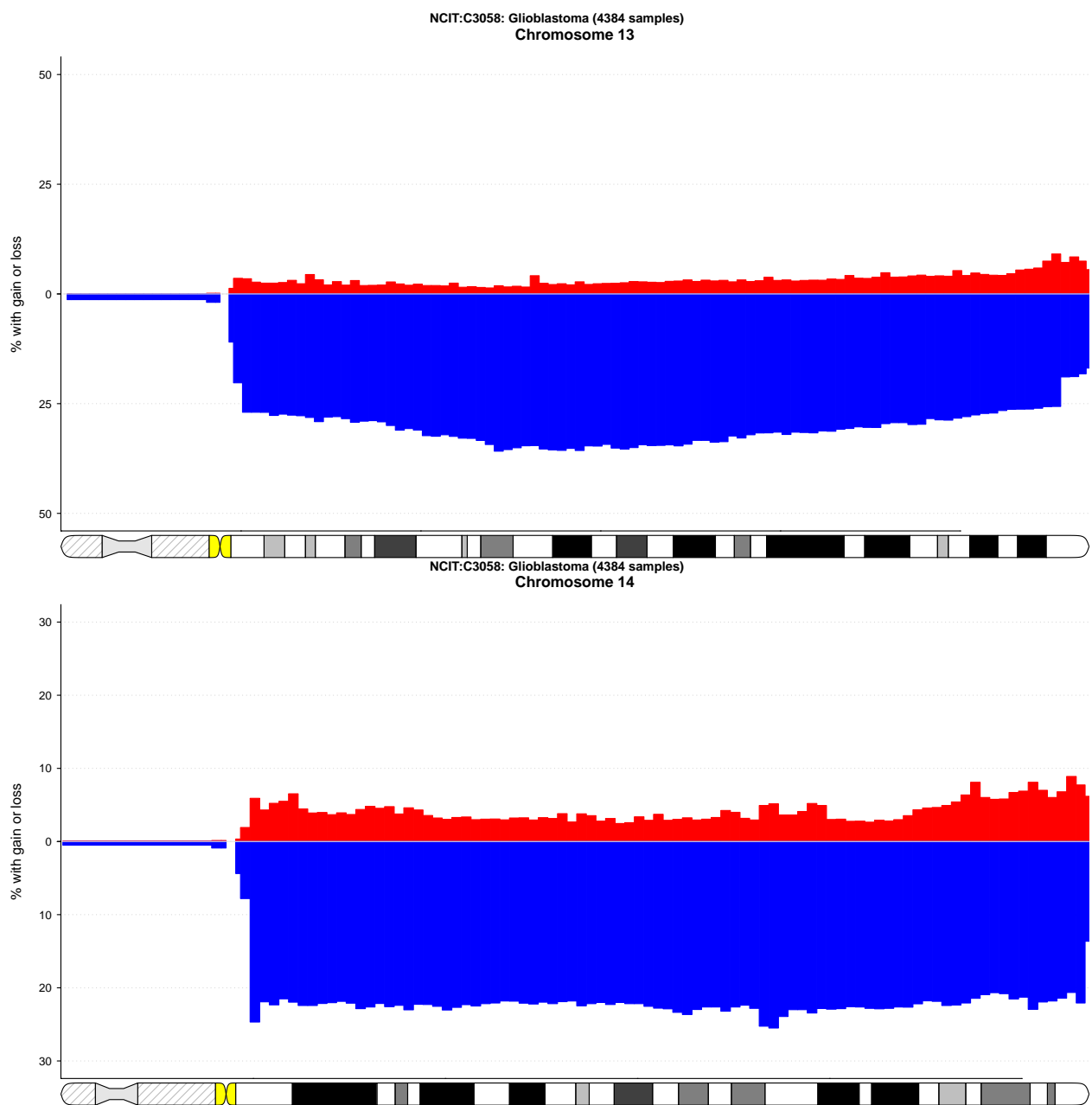
NCIT:C3058: Glioblastoma (4384 samples)
Chromosome 6

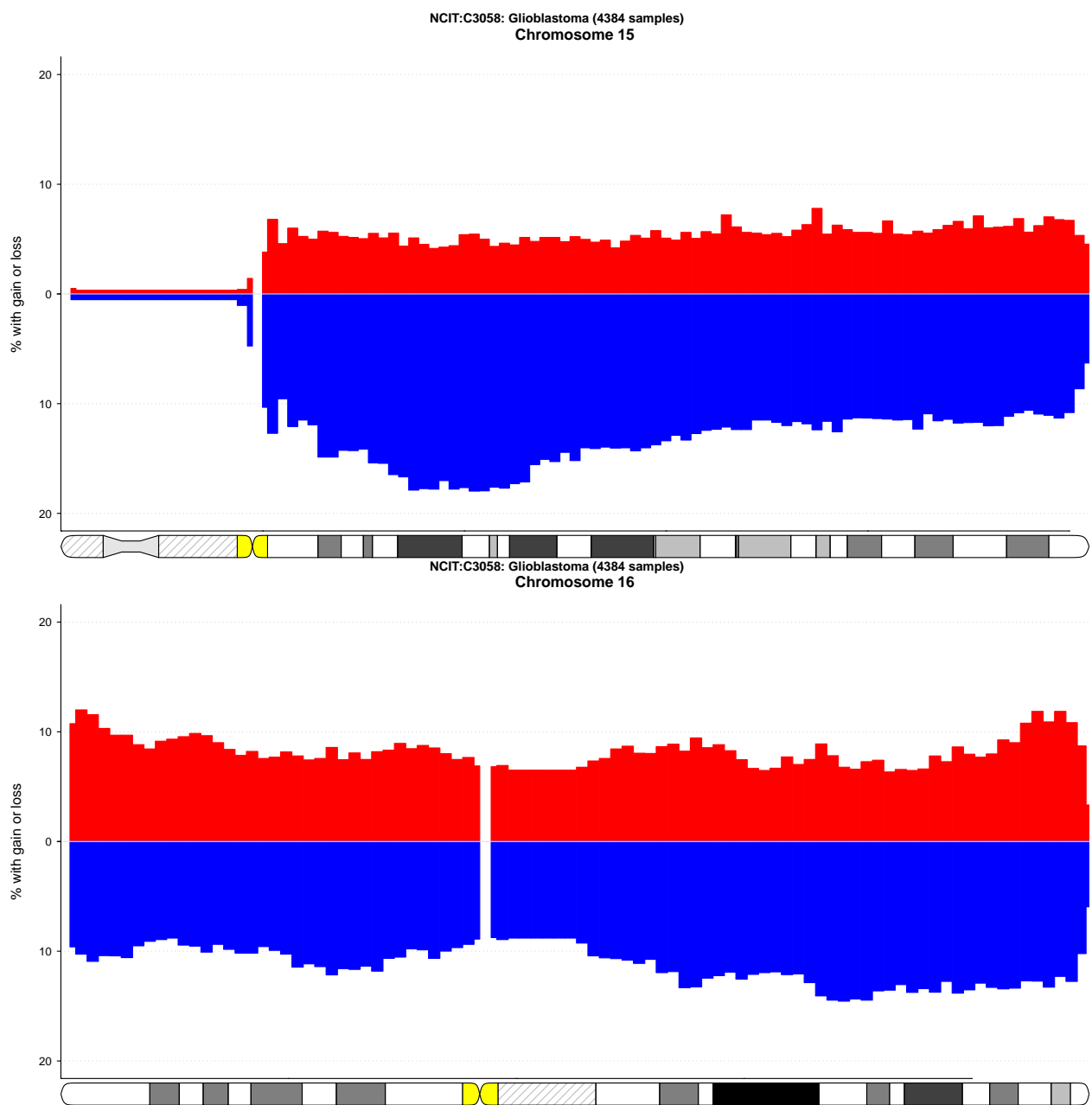


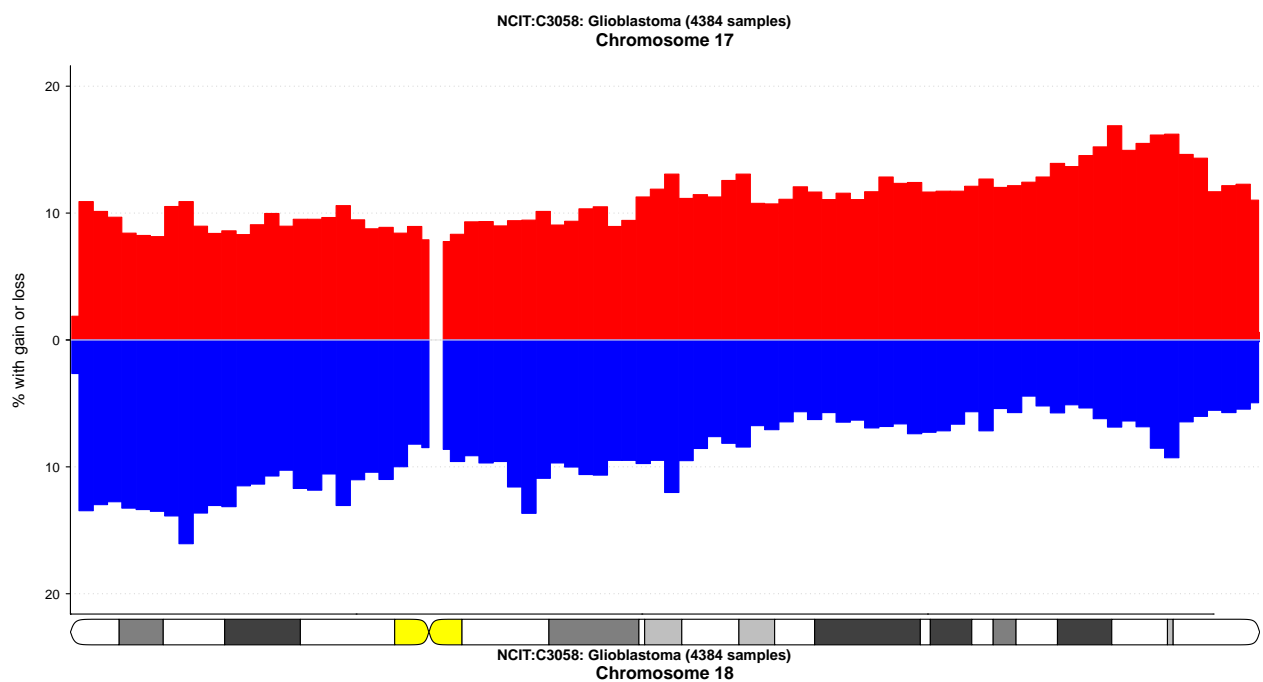




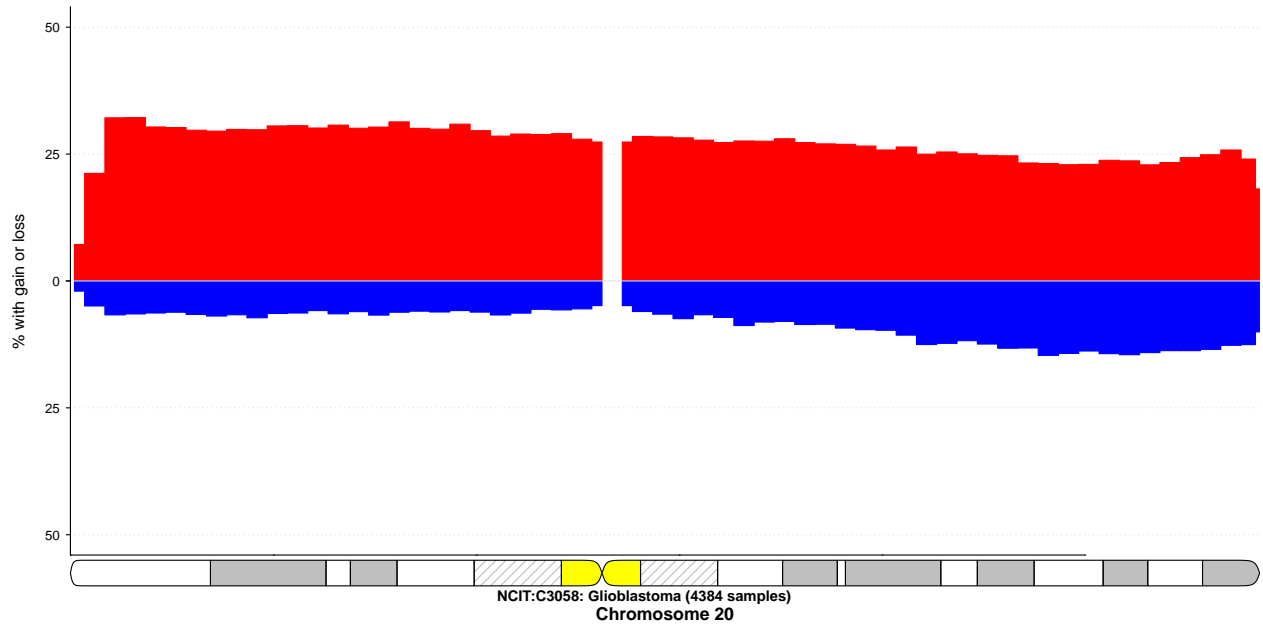






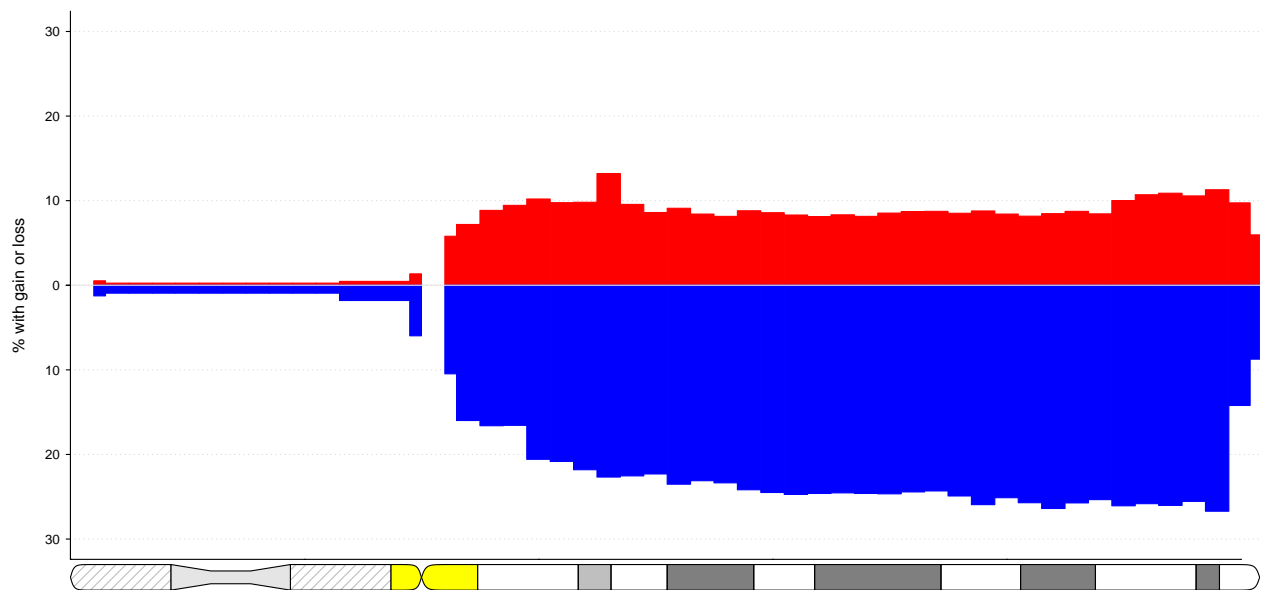
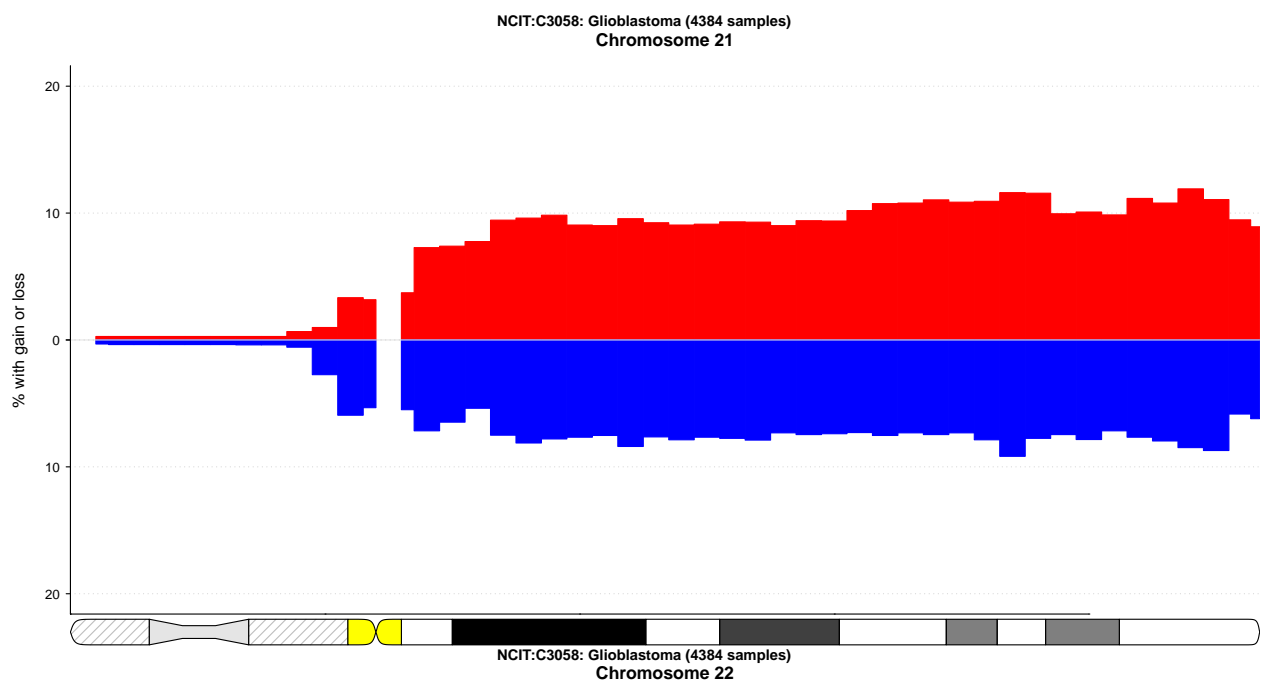


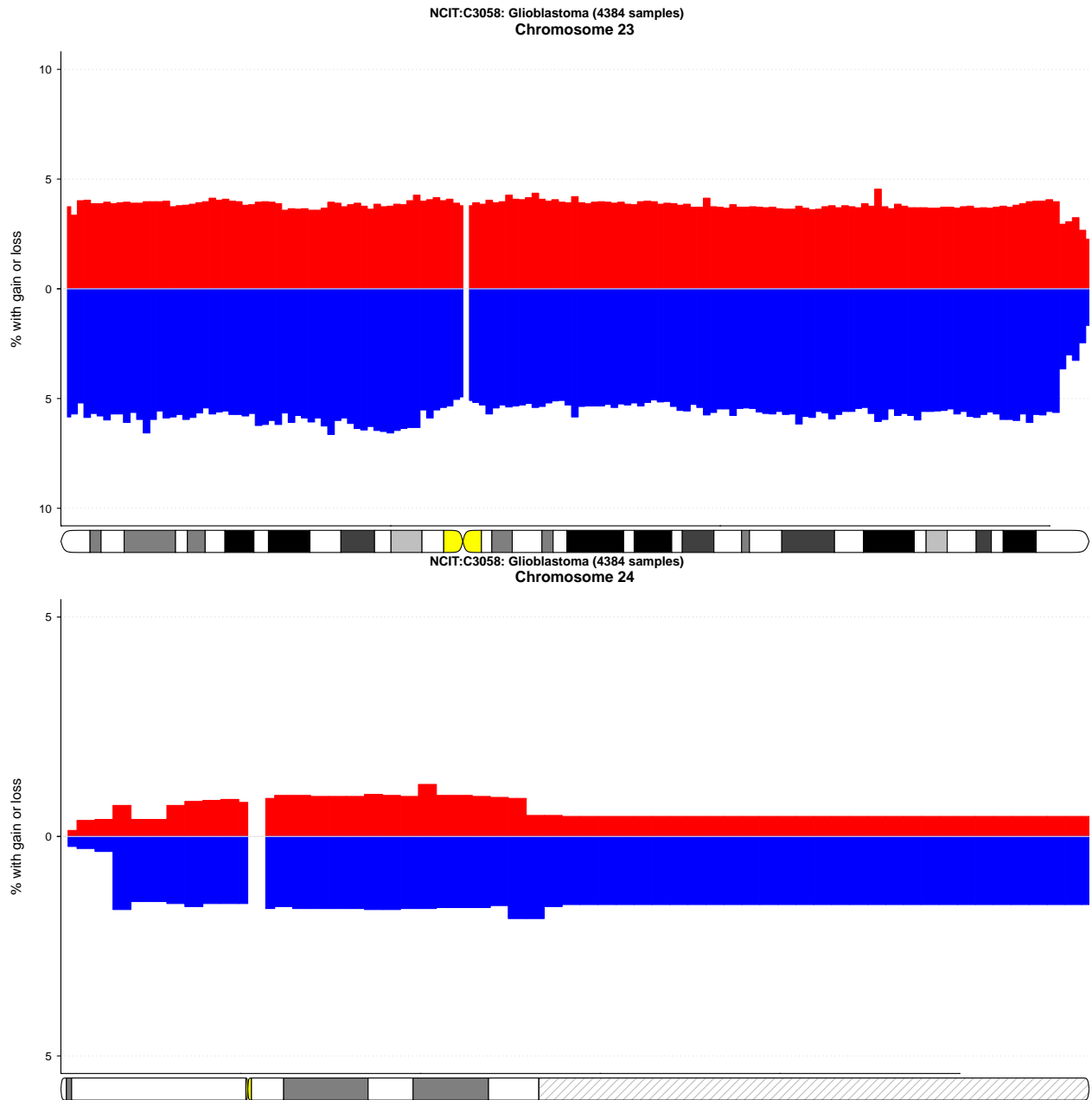
NCIT:C3058: Glioblastoma (4384 samples)
Chromosome 19



NCIT:C3058: Glioblastoma (4384 samples)
Chromosome 20







Step5: Analyse the data

According the plot, we can see frequenct gains on chromosome 7p, 8q, 20p,20q and frequenct losses on chromosome 4p,4q, 5q, 9p, 17p, 18q, 21q.

There is a literature where the findings are consistent with the majority of mine. Here is the paper-link.

A more detailed use case see this link.